



RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/502,984B

DATE: 07/03/2002

TIME: 12:12:59

Input Set : A:\A68126-1.app

Output Set: N:\CRF3\07032002\I502984B.raw

3 <110> APPLICANT: Luo, Peizhi
 5 <120> TITLE OF INVENTION: STRUCTURE-BASED SCREENING TECHNIQUES FOR DRUG DISCOVERY
 7 <130> FILE REFERENCE: A-68126-1/RFT/RMS/RMK
 9 <140> CURRENT APPLICATION NUMBER: 09/502,984B
 10 <141> CURRENT FILING DATE: 2000-02-11
 12 <150> PRIOR APPLICATION NUMBER: 60/120,009
 13 <151> PRIOR FILING DATE: 1999-02-11
 15 <150> PRIOR APPLICATION NUMBER: 60/131,674
 16 <151> PRIOR FILING DATE: 1999-04-29
 18 <160> NUMBER OF SEQ ID NOS: 37
 20 <170> SOFTWARE: PatentIn Ver. 2.1
 22 <210> SEQ ID NO: 1
 23 <211> LENGTH: 225
 24 <212> TYPE: PRT
 25 <213> ORGANISM: Homo sapiens
 27 <400> SEQUENCE: 1
 28 Ala Pro Pro Pro Asn Leu Pro Asp Pro Lys Phe Glu Ser Lys Ala Ala
 29 1 5 10 15
 31 Leu Leu Ala Ala Arg Gly Pro Glu Glu Leu Leu Cys Phe Thr Glu Arg
 32 20 25 30
 34 Leu Glu Asp Leu Val Cys Phe Trp Glu Glu Ala Ala Ser Ala Gly Val
 35 35 40 45
 37 Gly Pro Gly Asn Tyr Ser Phe Ser Tyr Gln Leu Glu Asp Glu Pro Trp
 38 50 55 60
 40 Lys Leu Cys Arg Leu His Gln Ala Pro Thr Ala Arg Gly Ala Val Arg
 41 65 70 75 80
 43 Phe Trp Cys Ser Leu Pro Thr Ala Asp Thr Ser Ser Phe Val Pro Leu
 44 85 90 95
 46 Glu Leu Arg Val Thr Ala Ala Ser Gly Ala Pro Arg Tyr His Arg Val
 47 100 105 110
 49 Ile His Ile Asn Glu Val Val Leu Leu Asp Ala Pro Val Gly Leu Val
 50 115 120 125
 52 Ala Arg Leu Ala Asp Glu Ser Gly His Val Val Leu Arg Trp Leu Pro
 53 130 135 140
 55 Pro Pro Glu Thr Pro Met Thr Ser His Ile Arg Tyr Glu Val Asp Val
 56 145 150 155 160
 58 Ser Ala Gly Asn Gly Ala Gly Ser Val Gln Arg Val Glu Ile Leu Glu
 59 165 170 175
 61 Gly Arg Thr Glu Cys Val Leu Ser Asn Leu Arg Gly Arg Thr Arg Tyr
 62 180 185 190
 64 Thr Phe Ala Val Arg Ala Arg Met Ala Glu Pro Ser Phe Gly Gly Phe
 65 195 200 205
 67 Trp Ser Ala Trp Ser Glu Pro Val Ser Leu Leu Thr Pro Ser Asp Leu

RAW SEQUENCE LISTING

DATE: 07/03/2002

PATENT APPLICATION: US/09/502,984B

TIME: 12:12:59

Input Set : A:\A68126-1.app

Output Set: N:\CRF3\07032002\I502984B.raw

```

68      210      215      220
70 Asp
71 225
74 <210> SEQ ID NO: 2
75 <211> LENGTH: 211
76 <212> TYPE: PRT
77 <213> ORGANISM: Homo sapiens
79 <400> SEQUENCE: 2
80 Lys Phe Glu Ser Lys Ala Ala Leu Leu Ala Ala Arg Gly Pro Glu Glu
81 1 5 10 15
83 Leu Leu Cys Phe Thr Glu Arg Leu Glu Asp Leu Val Cys Phe Trp Glu
84 20 25 30
86 Glu Ala Ala Ser Ala Gly Val Gly Pro Gly Asn Tyr Ser Phe Ser Tyr
87 35 40 45
89 Gln Leu Glu Asp Glu Pro Trp Lys Leu Cys Arg Leu His Gln Ala Pro
90 50 55 60
92 Thr Ala Arg Gly Ala Val Arg Phe Trp Cys Ser Leu Pro Thr Ala Asp
93 65 70 75 80
95 Thr Ser Ser Phe Val Pro Leu Glu Leu Arg Val Thr Ala Ala Ser Gly
96 85 90 95
98 Ala Pro Arg Tyr His Arg Val Ile His Ile Asn Glu Val Val Leu Leu
99 100 105 110
101 Asp Ala Pro Val Gly Leu Val Ala Arg Leu Ala Asp Glu Ser Gly His
102 115 120 125
104 Val Val Leu Arg Trp Leu Pro Pro Pro Glu Thr Pro Met Thr Ser His
105 130 135 140
107 Ile Arg Tyr Glu Val Asp Val Ser Ala Gly Asn Gly Ala Gly Ser Val
108 145 150 155 160
110 Gln Arg Val Glu Ile Leu Glu Gly Arg Thr Glu Cys Val Leu Ser Asn
111 165 170 175
113 Leu Arg Gly Arg Thr Arg Tyr Thr Phe Ala Val Arg Ala Arg Met Ala
114 180 185 190
116 Glu Pro Ser Phe Gly Gly Phe Trp Ser Ala Trp Ser Glu Pro Val Ser
117 195 200 205
119 Leu Leu Thr
120 210
123 <210> SEQ ID NO: 3
124 <211> LENGTH: 212
125 <212> TYPE: PRT
126 <213> ORGANISM: Artificial Sequence
128 <220> FEATURE:
129 <223> OTHER INFORMATION: Description of Artificial Sequence: SYNTHETIC
131 <400> SEQUENCE: 3
132 Lys Phe Glu Ser Lys Ala Ala Leu Leu Ala Ala Arg Gly Pro Glu Glu
133 1 5 10 15
135 Leu Leu Cys Phe Thr Glu Arg Leu Glu Asp Leu Val Cys Phe Trp Glu
136 20 25 30
138 Glu Ala Ala Ser Ala Gly Val Gly Pro Gly Asn Phe Ser Phe Ser Phe
139 35 40 45

```

RAW SEQUENCE LISTING

DATE: 07/03/2002

PATENT APPLICATION: US/09/502,984B

TIME: 12:12:59

Input Set : A:\A68126-1.app

Output Set: N:\CRF3\07032002\I502984B.raw

```

141 Gln Leu Glu Asp Glu Pro Trp Lys Leu Cys Arg Leu His Gln Ala Pro
142      50                      55                      60
144 Thr Ala Arg Gly Ala Ile Arg Ile Phe Trp Cys Ser Leu Pro Thr Ala
145 65                      70                      75                      80
147 Asp Thr Ser Ser Phe Val Pro Leu Glu Leu Arg Val Thr Ala Ala Ser
148                      85                      90                      95
150 Gly Ala Pro Arg Phe His Arg Val Ile His Ile Asn Glu Val Val Leu
151                      100                      105                      110
153 Leu Asp Ala Pro Val Gly Leu Val Ala Arg Leu Ala Asp Glu Ser Gly
154                      115                      120                      125
156 His Val Val Leu Arg Trp Leu Pro Pro Pro Glu Thr Pro Met Thr Ser
157                      130                      135                      140
159 His Ile Arg Tyr Glu Val Asp Val Ser Ala Gly Asn Gly Ala Gly Ser
160 145                      150                      155                      160
162 Val Gln Arg Val Glu Ile Leu Glu Gly Arg Thr Glu Cys Val Leu Ser
163                      165                      170                      175
165 Asn Leu Arg Gly Arg Thr Arg Tyr Thr Phe Ala Val Arg Ala Arg Met
166                      180                      185                      190
168 Ala Glu Pro Ser Phe Gly Gly Phe Trp Ser Ala Trp Ser Glu Pro Val
169                      195                      200                      205
171 Ser Leu Leu Thr
172      210

```

175 <210> SEQ ID NO: 4

176 <211> LENGTH: 211

177 <212> TYPE: PRT

178 <213> ORGANISM: Artificial Sequence

180 <220> FEATURE:

181 <223> OTHER INFORMATION: Description of Artificial Sequence: SYNTHETIC

183 <400> SEQUENCE: 4

```

184 Lys Phe Glu Ser Lys Ala Ala Leu Leu Ala Ala Arg Gly Pro Glu Glu
185 1      5                      10                      15
187 Leu Leu Cys Phe Thr Glu Arg Leu Glu Asp Leu Val Cys Phe Phe Glu
188                      20                      25                      30
190 Glu Ala Ala Ser Ala Gly Val Gly Pro Gly Asn Phe Ser Phe Ser Phe
191                      35                      40                      45
193 Gln Leu Glu Asp Glu Pro Trp Lys Leu Cys Arg Leu His Gln Ala Pro
194      50                      55                      60
196 Thr Ala Arg Gly Ala Ile Arg Phe Trp Cys Ser Leu Pro Thr Ala Asp
197 65                      70                      75                      80
199 Thr Ser Ser Phe Val Pro Leu Glu Leu Arg Leu Thr Ala Ala Ser Gly
200                      85                      90                      95
202 Ala Pro Arg Phe His Arg Val Ile His Ile Asn Glu Val Val Leu Leu
203                      100                      105                      110
205 Asp Ala Pro Val Gly Leu Val Ala Arg Leu Ala Asp Glu Ser Gly His
206                      115                      120                      125
208 Val Val Leu Arg Trp Leu Pro Pro Pro Glu Thr Pro Met Thr Ser His
209                      130                      135                      140
211 Ile Arg Tyr Glu Val Asp Val Ser Ala Gly Asn Gly Ala Gly Ser Val
212 145                      150                      155                      160

```

RAW SEQUENCE LISTING

DATE: 07/03/2002

PATENT APPLICATION: US/09/502,984B

TIME: 12:12:59

Input Set : A:\A68126-1.app

Output Set: N:\CRF3\07032002\I502984B.raw

```

214 Gln Arg Val Glu Ile Leu Glu Gly Arg Thr Glu Cys Val Leu Ser Asn
215           165           170           175
217 Leu Arg Gly Arg Thr Arg Tyr Thr Phe Ala Val Arg Ala Arg Met Ala
218           180           185           190
220 Glu Pro Ser Phe Gly Gly Phe Trp Ser Ala Trp Ser Glu Pro Val Ser
221           195           200           205
223 Leu Leu Thr
224       210
227 <210> SEQ ID NO: 5
228 <211> LENGTH: 211
229 <212> TYPE: PRT
230 <213> ORGANISM: Artificial Sequence
232 <220> FEATURE:
233 <223> OTHER INFORMATION: Description of Artificial Sequence: SYNTHETIC
235 <400> SEQUENCE: 5
236 Lys Phe Glu Ser Lys Ala Ala Leu Leu Ala Ala Arg Gly Pro Glu Glu
237   1           5           10           15
239 Leu Leu Cys Phe Thr Glu Arg Leu Glu Asp Leu Val Cys Phe Trp Glu
240           20           25           30
242 Glu Ala Ala Ser Ala Gly Val Gly Pro Gly Asn Tyr Ser Phe Ser Tyr
243           35           40           45
245 Gln Leu Glu Asp Glu Pro Trp Lys Leu Cys Arg Leu His Gln Ala Pro
246           50           55           60
248 Thr Ala Arg Gly Ala Val Arg Phe Trp Cys Ser Leu Pro Thr Ala Asp
249   65           70           75           80
251 Thr Ser Ser Phe Val Pro Leu Glu Leu Arg Val Thr Ala Ala Ser Gly
252           85           90           95
254 Ala Pro Arg Tyr His Arg Val Ile His Ile Asn Glu Val Val Leu Leu
255           100          105          110
257 Asp Ala Pro Val Gly Leu Val Ala Arg Leu Ala Asp Glu Ser Gly His
258           115          120          125
260 Val Val Ile Arg Trp Leu Pro Pro Pro Glu Thr Pro Met Thr Ser His
261           130          135          140
263 Ile Arg Phe Glu Leu Asp Ile Ser Ala Gly Asn Gly Ala Gly Ser Val
264 145           150           155           160
266 Gln Arg Val Glu Leu Leu Glu Gly Arg Thr Glu Cys Val Leu Ser Asn
267           165           170           175
269 Leu Arg Gly Arg Thr Arg Ile Thr Ile Ala Val Arg Ala Arg Met Ala
270           180           185           190
272 Glu Pro Ser Phe Gly Gly Phe Trp Ser Ala Trp Ser Glu Pro Val Ser
273           195           200           205
275 Leu Leu Thr
276       210
279 <210> SEQ ID NO: 6
280 <211> LENGTH: 211
281 <212> TYPE: PRT
282 <213> ORGANISM: Artificial Sequence
284 <220> FEATURE:
285 <223> OTHER INFORMATION: Description of Artificial Sequence: SYNTHETIC

```

RAW SEQUENCE LISTING

DATE: 07/03/2002

PATENT APPLICATION: US/09/502,984B

TIME: 12:12:59

Input Set : A:\A68126-1.app

Output Set: N:\CRF3\07032002\I502984B.raw

287 <400> SEQUENCE: 6

```

288 Lys Phe Glu Ser Lys Ala Ala Leu Leu Ala Ala Arg Gly Pro Glu Glu
289   1           5           10           15
291 Leu Leu Cys Phe Thr Glu Arg Leu Glu Asp Leu Val Cys Phe Phe Glu
292           20           25           30
294 Glu Ala Ala Ser Ala Gly Val Gly Pro Gly Asn Phe Ser Phe Ser Phe
295           35           40           45
297 Gln Leu Glu Asp Glu Pro Trp Lys Leu Cys Arg Leu His Gln Ala Pro
298           50           55           60
300 Thr Ala Arg Gly Ala Ile Arg Phe Trp Cys Ser Leu Pro Thr Ala Asp
301  65           70           75           80
303 Thr Ser Ser Phe Val Pro Leu Glu Leu Arg Leu Thr Ala Ala Ser Gly
304           85           90           95
306 Ala Pro Arg Phe His Arg Val Ile His Ile Asn Glu Val Val Leu Leu
307           100          105          110
309 Asp Ala Pro Val Gly Leu Val Ala Arg Leu Ala Asp Glu Ser Gly His
310           115          120          125
312 Val Val Ile Arg Trp Leu Pro Pro Pro Glu Thr Pro Met Thr Ser His
313           130          135          140
315 Ile Arg Phe Glu Leu Asp Ile Ser Ala Gly Asn Gly Ala Gly Ser Val
316 145          150          155          160
318 Gln Arg Val Glu Leu Leu Glu Gly Arg Thr Glu Cys Val Leu Ser Asn
319           165          170          175
321 Leu Arg Gly Arg Thr Arg Ile Thr Ile Ala Val Arg Ala Arg Met Ala
322           180          185          190
324 Glu Pro Ser Phe Gly Gly Phe Trp Ser Ala Trp Ser Glu Pro Val Ser
325           195          200          205
327 Leu Leu Thr
328           210

```

331 <210> SEQ ID NO: 7

332 <211> LENGTH: 211

333 <212> TYPE: PRT

334 <213> ORGANISM: Artificial Sequence

336 <220> FEATURE:

337 <223> OTHER INFORMATION: Description of Artificial Sequence: SYNTHETIC

339 <400> SEQUENCE: 7

```

340 Lys Phe Glu Ser Lys Ala Ala Leu Leu Ala Ala Arg Gly Pro Glu Glu
341   1           5           10           15
343 Leu Leu Cys Phe Thr Glu Arg Leu Glu Asp Leu Val Cys Phe Trp Glu
344           20           25           30
346 Glu Ala Ala Ser Ala Gly Val Gly Pro Gly Asn Phe Ser Phe Ser Phe
347           35           40           45
349 Gln Leu Glu Asp Glu Pro Trp Lys Leu Cys Arg Leu His Gln Ala Pro
350           50           55           60
352 Thr Ala Arg Gly Ala Ile Arg Phe Trp Cys Ser Leu Pro Thr Ala Asp
353  65           70           75           80
355 Thr Ser Ser Phe Val Pro Leu Glu Leu Arg Val Thr Ala Ala Ser Gly
356           85           90           95
358 Ala Pro Arg Phe His Arg Val Ile His Ile Asn Glu Val Val Leu Leu

```

RAW SEQUENCE LISTING ERROR SUMMARY DATE: 07/03/2002
PATENT APPLICATION: US/09/502,984B TIME: 12:13:00

Input Set : A:\A68126-1.app
Output Set: N:\CRF3\07032002\I502984B.raw

Please Note:

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:30; Xaa Pos. 3

VERIFICATION SUMMARY

DATE: 07/03/2002

PATENT APPLICATION: US/09/502,984B

TIME: 12:13:00

Input Set : A:\A68126-1.app

Output Set: N:\CRF3\07032002\I502984B.raw

L:1542 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:30 after pos.:0